DOCUMENT RESUME

ED 161 512 PS 009 913

AUTHOR Takanishi, Ruby
TITLE Evaluation of Early Childhood Programs: Toward a Developmental Perspective.
INSTITUTION ERIC Clearinghouse on Early Childhood Education, Urbana, Ill.
SPONS AGENCY National Academy of Education, Syracuse, N.Y.; National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 78
NOTE 51p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS Behavior Theories; *Child Development; *Developmental Psychology; *Early Childhood Education; Educational Theories; *Models; Program Effectiveness; *Program Evaluation

ABSTRACT

This paper presents a rationale for taking a developmental perspective towards evaluation of early childhood programs and contrasts two models of development (organismic and mechanistic) with regard to evaluation and developmental issues. Examination of theoretical differences between the organismic model (which assumes an active organism) and mechanistic model (which assumes a reactive organism) suggests that a developmental perspective: (1) implies that behavioral observations must be analyzed with respect to their meaning in context; (2) raises the question of the meaningfulness of educational testing situations; (3) is important in the conceptualization of development in early childhood programs; and (4) leads to a shift away from a preoccupation with criterion performance to a concern with the processes underlying performance. A developmental perspective also emphasizes: changes in cognitive structures rather than behavioral change per se; discontinuity rather than continuity of change; reciprocal causality between organism and environment rather than unidirectional causality of environment on organism; and organized complexity rather than linear causality in change. The implications of these model features for assessing program effectiveness are discussed. The rationale presented for working toward a developmental perspective in evaluation highlights the involvement of participants in understanding program development, adult (staff and parent) development, and child development. (JMB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

EVALUATION OF EARLY CHILDHOOD PROGRAMS: TOWARD A DEVELOPMENTAL PERSPECTIVE[1]

Ruby Takanishi

University of California, Los Angeles

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ruby Takanishi

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED161512

PS009913

The widespread evaluation of social action and educational programs is barely a decade old (Ross & Cronbach, 1976). Much of this evaluation activity has centered on programs serving low-income and minority group children and their families. The mandate for the evaluation of these early childhood programs took the fields of developmental psychology and early childhood education by surprise. Few individuals existed in either field in the mid-'60s who were knowledgeable and trained in evaluation, which itself lacked a theory unique to its special problems (Guba, 1969). In this vacuum, evaluation became associated with the measurement of child outcomes which could be reliably assessed. One result was that certain valuative, epistemological, and conceptual assumptions related to research with young, developing children were overlooked.

For a number of reasons, the evaluation of programs serving young children and their families continues to be an important social and scientific issue. The most important reason for continued controversy is the diversity of strongly held beliefs regarding the effectiveness of these programs among diverse constituencies (Katz, 1975; Zigler, Note 10). This paper acknowledges the inherently political nature of evaluation activities (Datta, 1976; House, 1973), but focuses as a task on a developmental perspective to evaluation and how this perspective can inform future evaluation studies of early childhood programs.

The major points of this paper can be summarized.

1. The evaluator's model of development, whether explicit or implicit, is reflected in his design and doing of evaluation. Models of development influence decisions about what are considered meaningful problems or questions to pose in an evaluation, what methods of data collection and analysis will be used, how the data will be interpreted, and what implications for policy are drawn. Progress in evaluating early childhood programs should not be only associated with increasingly sophisticated research designs and advanced statistical techniques. Fundamental questions about the nature of human development as well as the purposes of evaluation are also involved.

2. The evaluation of early childhood programs, like most evaluation studies, is dominated by psychometric, experimental, and behavioristic models (Eisner, 1977; Guba, 1969; House, 1976; Levine, 1974; Patton, Note 4). Program outcomes have been limited to standardized measures which can be reliably assessed, e.g., IQ and achievement, with less attention to the meaningfulness of what is being measured (Messick, 1975). Within the dominant evaluation model, change is defined quantitatively as the acquisition of more pieces of information, knowledge, and experiences without attention to the structures underlying observed changes. A child's behavior or performance is assessed to determine his position in relation to single standards, e.g., intelligence tests (Riegel, 1972).

3. Evaluation lacks a perspective which is grounded in the nature of developmental change and in which means-ends relations and their transformations are central. I will argue that there is an incongruency between current evaluation strategies and the phenomena under study.

Within a developmental perspective, change is considered to be qualitative as well as quantitative: "New properties emerge, irreducible to lower levels and, therefore, qualitatively different from them" (Overton & Reese, 1973, p. 70). There are multiple influences on development which result from a complex transaction between internal and external forces. The nature of developmental change is considered to be dynamic and differentiated with multiple outcomes as well as multiple pathways to similar behavior. A developmental perspective encourages multigenerational and multicultural standards (Riegel, 1972).

4. The evaluation of early childhood programs would benefit from a developmental perspective to guide its activities. The purposes of evaluation will then be changed to focus on program, adult, and child development. This paper represents some steps toward a developmental perspective to evaluation and is far from being complete. The developmental perspective must be considered, because of the importance of evolving evaluation strategies that are uniquely suited to early childhood programs, and that also share in emergent models in the general field of evaluation itself (Hamilton, et al., 1977; Willis, 1978).

Models of Development

Katz (1975), in her discussion of early childhood programs and ideological disputes, observed that in the formal research and evaluation literature the exchange of divergent views concerning what young children need and how these needs should be satisfied are typically couched in the language of theory, methodology, and evaluation. She argued that the conflicts are not theoretical ones, but ideological ones related to strongly held conceptions of childhood, development, and the good life, which cannot be settled by appeals to evidence. Several other writers

in child development have pointed out that a major problem in the evaluation of early childhood programs is the failure to relate evaluation strategies to multiple views of learning and development (Almy, 1975; Kamii & Elliott, 1971; Kohlberg & Mayer, 1972; Murphy, 1973; Sigel, 1972).[2]

For the purposes of this paper, it is argued that the evaluator's idea or model of development with its related values is reflected in his design and doing of evaluations. Models create lenses from which certain phenomena are seen and others excluded (Petrie, 1972). They determine what are considered meaningful questions for an evaluation, what methods of data collection and analysis will be used, how the data will be interpreted, and what implications for policy are drawn. Evaluators are socialized to certain values and models of development as part of their training and to a shared sense of what is acceptable to their colleagues or reference groups. Evaluation involves evaluators-as-persons involved in networks of scholarly, social, and political relationships which, in turn, affect their evaluation designs (Sjoberg, 1975).

The evaluations of Sesame Street illustrate the fact that two groups of investigators with different value perspectives asked different questions and derived different conclusions. Ball and Bogatz (1970) accepted and used the curriculum goals of the program and found that Sesame Street met the goal of stimulating the growth of children who watched the show more than comparable groups that did not. Cook and his associates (1975) interpreted the goal of Sesame Street to be a decrease the gap between low- and middle-income children in terms of learning, and concluded that the program was not successful in narrowing the gap between the two groups.

In agreement with Kuhn's (1962) notion of science as ideology, a model cannot be judged true or false, but only more or less useful as a model from which to view reality: ". . . since no paradigm ever solves all the problems it defines and since no two paradigms leave all the same problems unsolved, paradigm debates always involve the question: Which problem is it more significant to have solved? Like the issue of competing standards, that question of values can be answered only in terms of criteria that lie outside of normal science altogether, and it is that recourse to external criteria that most obviously makes paradigm debates revolutionary" (Kuhn, 1962, pp. 108-109).

The argument for a developmental perspective to the evaluation of early childhood programs is based in the belief that the most significant problem for evaluation of programs for young children is how programs affect their and other participants' (parents and staff) full development as human beings. Central to this argument is a consideration of alternative models of conceptualizing development and developmental change. In order to organize the following argument for a developmental perspective to the evaluation of early childhood programs, two models of development and their corollary assumptions as described by Reese and Overton (1970; Overton & Reese, 1973; see also Looft, 1973) will be utilized as a foundation.

Reese and Overton (1970) identified two models of development in psychology: the organismic (active organism) and mechanistic (reactive organism) world views. Their position, shared by this author, is that these models reflect different ways of looking at phenomena and are incompatible in their implications (see above, Kuhn, 1962). This position does not mean that the use of one model precludes the other. At different

stages in an evaluation, one of the models may be more useful in answering the questions posed. However, the integration of the two into a common model is not possible. Both models of development have evolved over time, and often are not represented in their "pure" forms. These refinements will be noted in the course of the discussion.

The mechanistic model. This model is based on the metaphor of the machine. Component parts and their relationships form the reality to which complex behavior can be reduced and expressed in quantitative and functional relationships often called laws. The meanings of behavior remain constant over time. Prediction is possible and important since the knowledge of the machine at one point in time allows for inferences about the state of the machine at another point in time, given knowledge of past and potential efficient or material causes.

Development is defined in terms of observable behavior, and theories of development are extensions of learning theories. The epistemological position is that of naive realism, in which the knower plays no role in the known and is assumed to apprehend the world in a standard way. Thus, development can be assessed by how individuals measure up to single standards. Change occurs over time, and does not result from changes in the structures of the organism.

The organismic model. This model is based on the metaphor of person as actively influencing and shaping experience. The epistemological position is that of constructivism (Magoon, 1977). The knower, in this case the child, actively constructs reality (Harre, 1974; Kamii, 1975). What is known is the product of the interaction between the active organism and the environment. Knowledge is gained only as the knower constructs the world.

Development is defined as change in the organization of a structure which cannot be directly observed and must be inferred from behavior. The relationship between structure or processes and functions or purposes is central. The organismic model assumes qualitative change in that the meaning of behaviors change through the course of ontogenesis. Since change is the result of formal causes rather than efficient causes (although efficient causes may inhibit or facilitate change), the possibility of a strictly predictive and quantifiable world is precluded.

The organismic model of development has rarely been represented in evaluations of early childhood programs, but is pervasive in Piagetian preschool curricula (Kamii, 1973; 1975). In the sense that this model is closely related to what have been called qualitative (Hamilton, et al., 1977; Willis, 1978) and constructivist approaches to educational research and evaluation (Magoon, 1977), these approaches are not well represented in evaluation in general. Educational evaluation is dominated by a model which focuses on products (behavior and achievement) over processes, and on quantitative change rather than qualitative change (Eisner, 1977; Levine, 1974; Partlett & Hamilton, 1976; Patton, Note 4).

Corollary Model Issues

Reese and Overton (1970; Overton & Reese, 1973) identify corollary model issues which affect the analysis and understanding of development. These issues are considered to be model-like or pre-theoretical and not open to empirical test. These issues will be described with their implications for the evaluation of early childhood programs.

Elementarism versus holism. The mechanistic model holds that the whole is predictable from its parts, that physically identical elements have the same "meaning." In contrast, the organismic model considers

the person as an organized totality in which the parts derive meaning from the whole. Reese and Overton (1970) state: "Rather than assessing behavior in terms of material identity, this (holistic) assumption directs assessment in terms of the function of the behavior in the whole or context in which it is embedded, that is, according to the function or ends, or goals of the organism or part processes which are being investigated" (p. 137).

This statement has several implications for the evaluation of early childhood programs, particularly in the measurement realm. The idea that behaviors that are physically identical do not have the same meaning places the organismic model squarely in the qualitative-phenomenological (Merleau-Ponty, 1962), and ethnographic tradition (Wilson, 1977). The phenomenological tradition asserts that human behavior cannot be understood without understanding the subject's interpretation of the situation, a position consistent with Piagetians (Kamii, 1975). Furthermore, there are multiple interpretations of any given situation depending upon the participants and their perspectives, and the past and current social context of the observed behavior. The phenomenological tradition also intersects with the psychodynamic one since both are concerned with latent and underlying meanings in participants' experiences. (For a psychodynamic perspective on curriculum evaluation, see Tyler, 1978).

The use of observational systems in evaluations to obtain information on classroom processes and behavior is typically concerned with behavior as "facts" rather than with the meaning of the observations in relation to the child, teacher, and the social milieu of the classroom. The need to examine further how children interpret and construct their experiences in early education classrooms is highlighted by Karlson's (cited in

Shapiro, 1973) observations that within a Montessori preschool, each child created his own curriculum. Shapiro (1973) also observed that children do not have the same experiences in the same physical setting such as the classroom. Thus observational data on the "implementation" of a certain curriculum does not mean that all individuals within the setting had the same experiences of that curriculum. The mechanistic model's assumption that physically identical elements have similar meaning is thus challenged. Assessment of classroom environments and curricula must be redirected in terms of how different individuals interpret classroom environments and curricula (Tyler, 1978; Mehan, et al., Note 3; Takanishi & Spitzer, Note 8). Similarly, observations of teacher and child behavior must be collected and analyzed in terms of their meaning for the participants (Magoon, 1977; Wilson, 1977).[3]

Moving from observations of classroom environments and behavior, the holistic-elementaristic distinction also has important implications for assessing child outcome variables in evaluations. Messick (1975), in reviewing the status of constructs, meanings, and values in educational research and evaluation, points out that educational measurement has been concerned mainly with predictive and content validity, and has neglected an important form of validity, namely construct validity. He emphasizes that "the meaning of the measure must be also pondered in order to evaluate responsibly the possible consequences of its proposed use" (p. 956).

Since intelligence, achievement, and other standardized tests have been heavily used in assessing child outcomes in evaluations, the focus of the discussion will be on these tests as meaningful measures. The controversies surrounding intelligence testing are well-known, focusing

on genetic and environmental factors (Jensen, 1969). Attention has also centered on the testing situation as a special setting constraining the child's behavior (Cazden, 1970; Shapiro, 1973; Sigel, 1973). Others have challenged the predictive validity of intelligence and achievement tests with reference to success within and outside the educational system (Kohlberg & Mayer, 1973; McClelland, 1973; Sigel, 1972). These arguments and positions are relatively well-known and will not be elaborated further.

However, critiques of the testing situation have moved to another level, namely, inquiry into the meaning of standardized assessment situations to children[4] (Cicourel, et al., 1974; Mehan, 1978; Murphy, 1973; Tyler, 1978). Cicourel and his associates conducted a number of studies related to the social context of intelligence testing and children's interpretative competence in the testing situation. These workers found that answers judged to be "wrong" may result because the child does not share the test constructors' constructions of correctness (for a similar argument, see Messick, 1975). Validation of Cicourel, et al. (1974) findings can be found in studies conducted in the Piagetian tradition (Kamii, 1973; 1975). The standardized testing situation is based on the assumption all children will approach the items in the same way. However, studies on the meaning of standardized test items indicate that a child's "incorrect" answer may result from a different interpretation of the testing materials (Mehan, 1978). Students with incorrect answers were often found to be performing the very cognitive operation being tested by the questions. The assumption that low scores necessarily reflect lack of ability in the child, particularly when the child is a member of a minority group: "It is necessary to examine the structure

of the child's accounting practices and reasoning processes in order to draw valid inferences about his competence" (Cicourel, et al., 1974, p. 5).

Another significant challenge to the assumptions of standardized testing comes from Mehan's (1978) careful studies of videotaped testing situations which show that testing results are jointly produced through the child and tester's interactions. Particularly in individually administered tests, testers emphasize certain words, praise correct answers, or cut off the child's questions. The child may ask questions of the tester which require him to respond when guidelines for administration caution against engaging in interaction with the child. Test results are thus socially negotiated in the interactions between child and tester.

To summarize, the implications of the elementarism-holism distinction can be found in the issue of the meaningfulness of behavioral observations and of educational testing situations. The distinction implies that observations cannot be dealt with solely at the level of facts, but must be analyzed with respect to their meaning-in-context. Observations of classroom life must be analyzed in terms of their meaning for the participants (Magoon, 1977). Educational testing situations must be seen as constructions of the test makers which incorporate their use of language and values, and these constructions may not be shared by those being assessed (Almy, 1975; Mehan, 1978; Messick, 1975). Thus, the validity of the testing situations and information obtained therein are challenged.

The elementarism-holism distinction is also important in relation to conceptualizations of development in early childhood programs. Within the last 15 years, there has been a diversification of the curriculum

with concommitant expansion of conceptions of children's development. A conception of the "child-as-a-whole," however, has the longest tradition. Almy (1975) describes this concept as both emphasizing the uniqueness of the individual child, and the organized whole of physical, mental, and social development. The concept necessitates the consideration of the impact of programs on all aspects of the child's development.

These ideas can be found in arguments for "comprehensive," multidimensional assessment in evaluation (Frank, 1968). A major obstacle to multidimensional assessment, however, is the lack of reliable and valid instruments in the social and emotional domains in particular. Likewise, the focus of instrument development has tended to be more developed in assessing what has been learned rather than how a person learns. As Zimiles (1977) has argued, a nomothetic, standardized approach to personality assessment required by large-scale evaluations that attempt comprehensiveness is currently not feasible, and he suggests an idiographic approach to personality assessment. Likewise, Shipman et al. (Note 6) found in their intensive case study examination of low-income, Black children that a given aspect of individual functioning must be evaluated in relation to other aspects and in the environmental conditions in which the child was living. She too argued for a multidimensional assessment of individuals and their environments in future studies, since for different children, different clusters of variables appear to be differentially important at different points in time (see also Sigel, 1972).

Structure-function versus antecedent-consequent. In the mechanistic model, analyses focus on antecedent-consequent relations or causes and effects. The organismic model focuses on the relationship between the

operation of structures and functions. In the former, the locus of the critical aspect of developmental dynamic is external; the organism changes in response to external forces. In the latter, the locus is internal to the organism, as structures within the organism change (Looft, 1973).

In discussing meaning in evaluation, Messick (1975) argued: "The major point is that in evaluating the efficacy of programs it is not sufficient simply to gauge the size of effects or to appraise input-output differences relative to costs. In addition, one should seek evidence about the nature and meaning of the processes that produced the effects, for an understanding of these instrumentalities is necessary for a full and proper judgment of the value of the outcome" (p. 963). A critical aspect of this distinction in the organismic model is that qualitatively different structures, processes, or modes of operation may result in the same behavioral achievements. The conception of multiple pathways to similar observed behavior is a key feature of a qualitative growth model (Riegel, 1972).

A related value assumption is that some structures are "higher order" than others, even though each may lead to the same behavior. In his critical remarks about the evaluation of preschool interventions, Glick (1968) pointed to the importance of making distinctions between behavioral achievement and the processes or structures underlying the achievement. He argued: "It is not enough to simply demonstrate that criterion performance (that is, achievements) increase with age or are changed by intervention. What is necessary in order to make any argument which is basic to developmental questions is to show that the processes underlying the achievement have in fact been shifted to a higher

developmental level" (p. 218). Glick suggested that the study of processes underlying behavioral achievement could focus on transferability of behavior and on analysis of patterns of responses in given situations.

One implication of the structure-function/antecedent-consequent distinction is a shift away from a preoccupation with criterion performance or outcomes to a concern with the processes underlying performance. This shift can also be characterized as a move from quantitative models to qualitative models of development, and [illegible]n with the transformational characteris ics of developmental change (Wohlhill, 1973).

The distinction between behavioral achievement and structures is further highlighted in examinations of the relationship between psychometric and Piagetian assessments of intelligence. DeVries and Kohlberg (1977) note that the psychometric conception of intelligence is based on the assumption that assessment can be based on the number of right answers the child gives relative to other children of the same chronological age, i.e., their position on the normal curve. In contrast, a Piagetian conception views intelligence in terms of an individual's place in a universal, invariant sequence of development, through which individuals pass at different rates. Assessment includes not only "right" answers, but the analysis of "wrong" answers and children's reasoning behind their answers.

Piaget (Evans, 1973, cited in DeVries & Kohlberg, 1977) has contended that intelligence tests measure performance and do not get at internal structures or operations. DeVries and Kohlberg (1977) studied the empirical relationship between psychometric and Piagetian tests of intelligence. Their results indicated that Piagetian, developmental stage tests measure something in common with and are also distinguishable

from psychometric tests of primary mental abilities. DeVries (1974) found that while Piagetian, IQ, and achievement tests overlapped to some degree, they each measured qualitatively different aspects of cognitive functioning.

These studies indicate that the distinction between performance and structures or processes is a useful one for the evaluation of early childhood programs. Performance is typically measured by IQ and achievement tests, and their predictive validity in terms of life success has been questioned (Kohlberg & Mayer, 1973; McClelland, 1973; Sigel, 1972). The failure to make a distinction between performance and structure is a critical one if the purposes of early childhood programs are <u>really</u> aimed at optimal human development, rather than increased school achievement (Kohlberg & Mayer, 1973).

<u>The Nature of Developmental Change</u>

The nature of developmental change and how that change occurs are assumptions (Overton & Reese, 1973; Reese & Overton, 1970) which can also be related to the evaluation of early childhood programs. These assumptions include whether change is conceptualized as behavioral versus structural, or continuous versus discontinuous. Additional assumptions are whether change occurs by unidirectionality causality versus reciprocal causality or linear causality versus organized complexity.

<u>The nature of behavioral versus structural change</u>. In the mechanistic model, it is behavior or responses that change over time. In the organismic model, inferred structures and functions change over time. In the former model, change is determined by efficient or material causes; in the latter, change occurs in the structure or organization and functions as development moves toward a goal or purpose.

Glick (1968) argued that the distinction between the performance of an act and the capability to perform that act has profound consequences for the assessment of children's development in programs. We cannot infer from changes in performance, such as an intelligence test, that the underlying cognitive structure or ability has changed. Zigler and Butterfield (1968) demonstrated that the Binet scores for low-income children varied by a mean of 10 points depending upon whether the child was tested under standard testing conditions or under conditions designed to make the child comfortable and thus to obtain his optimal level of performance. Zigler and Butterfield compared standard and optimal forms of IQ testing on children prior to entering a preschool program and seven months later. They found that while tested IQ showed a rise during the year, optimal IQ scores did not. Other researchers have also pointed to the importance of the distinction between performance in the standardized testing situation and the child's capacity to perform in other situations (Cazden, 1970; Cicourel, et al., 1974; McClelland, 1973; Murphy, 1973; Shapiro, 1973). Accordingly, the focus of attention moves from "gains" in scores to the underlying structures.

Continuous versus discontinuous change. Integrally related to the behavioral change-structural change dimension is whether developmental change is continuous or discontinuous. Mechanistic models are characterized by notions of continuous, additive, linear change. Organismic models assume discontinuity in development. (Looft (1973) notes that continuity is part of an organismic approach in a derivative sense.) Since change in the organization of parts results in an emergent system, change cannot be predicted from knowledge of the parts.

Unidirectional causality versus reciprocal causality. Reese and Overton (1973) distinguish between unidirectional causality in which the effect is dependent upon some external cause versus reciprocal causality in which both the environment and the organism affect each other in an ongoing manner.

Linear causality versus organized complexity. A mechanistic model assumes there is a linear relationship between a cause and an effect, individual causes are additive in their effects, and that causation is unidirectional. However, a developmental perspective takes as its explanation the idea of organized complexity, that is, changes in the organization of the parts. In light of reciprocal causality and organized complexity, explanation of developmental change is not possible in terms of efficient or material causes. What might facilitate development in one person may not operate similarly for another. The idea of organized complexity is also inclusive of cultural, socioeconomic, and regional differences in development.

Sigel (1972) described change as organized complexity resulting from preschool intervention programs. The child is composed of a variety of subsystems (perceptual, cognitive, emotional, etc.) whose relationships to each other vary over time. Change in one subsystem is related to changes in others. Sigel noted that even though development is overall a cumulative process, "the cumulative effect may express itself in various effects at different times" (p. 369). The distinctions regarding the nature of developmental change have profound implications for the evaluation of early education programs. Evaluation of early childhood programs, almost without exception, exhibit a unidirectional bias in which the program is viewed as the "treatment" which causes some change

in the child. This unidirectional bias is often reflected in discussions of the "predictive validity" of a program. The complexities involved in such prediction are highlighted by Shipman et al. (Note 6) who caution that "the prediction of a child's achievement from early indices of the home environment should not be interpreted to mean that these predictors necessarily determine the child's achievement. Families, children, and schools can and do change, with corresponding change in the nature of their interactions, and such change can be facilitative or harmful" (pp. 48-49; see Sigel (1972) for a similar argument).

If one takes a developmental perspective, there is much less concern with prediction of later behavior, given multiple, interacting, and conflicting influences on behavior. Based on her own longitudinal studies (Murphy, 1962) and those at the Institute of Human Development (Jones, Bayley, Macfarlane, & Honzik, 1971), Murphy (1973) stated, "To a large extent, each child's development is a mystery story whose outcome we cannot really predict" (p. 344). Jones et al. (1971), in their studies of the physical, mental, emotional, and social development of individuals, point to the difficulty in predicting in these areas, and stress the individuality and plasticity of growth patterns. In a recent publication based on the Berkeley Growth Study data, McCall, Eichorn, and Hogarty (1977), conceptualized change in mental development as reflecting periods of instability of individual differences and/or discontinuities in developmental function across age. The emphasis of their study was on locating and describing developmental change and transition, not continuity and stability.

The best, and perhaps only, example of the complexities involved in prediction from preschool intervention programs is the Shipman et al.

(Note 6) case studies of low-income Black children who were participating in the ETS-Head Start longitudinal study. Given the dynamic interrelations among physical, affective, social, and cognitive development, Shipman and her colleagues concluded it is extremely difficult to predict whether a child who is doing well in Head Start will continue to do so in his later elementary school years, or whether a child who is doing badly will continue to have problems in later life. The Shipman et al. conclusions are supported by those of other longitudinal studies (Murphy, 1962; Jones et al, 1971). Given the complexities involved, a developmental perspective to evaluation needs to be reoriented toward description and understanding (Strike, 1972) and toward concurrent, contemporary validity (McCall, 1977) rather than prediction. McCall's distinction between the continuity/discontinuity of developmental functions and the stability/instability of individual differences requires closer attention in longitudinal studies.

The evaluation of early childhood programs could benefit from a reciprocal causation perspective in which both children and their environments are viewed as changing over time (Riegel, 1972; Sameroff, 1975). S. White (Kilmer & Weinberg, 1974) argued for a conceptualization of the child as capable of making "multiple cognitive adaptations to (a variety of) contextual specific demands" (p. 61). A reciprocal causality approach is suggested by Shapiro's (1973) discussion of how children in early education programs "construct" their own curriculum within the program. Through the child's own activity and interpretation, a different "treatment" can be said to exist for different children in the program, resulting in a variety of outcomes for different children (Zimiles, 1977).

Implications for assessing program effectiveness. A developmental perspective to evaluation accepts the idea that there are serious limitations to proving the "effectiveness" of early childhood programs. Furthermore, this perspective rejects an efficient cause approach to studying program impact. There are other difficulties in proving that our programs are effective (McCall, 1977). Most of the evaluation studies are correlational and causality cannot be inferred. Even in those projects which are experimental and under more strict control and monitoring than the majority of programs, other limitations are present. There are multiple pathways to any observed outcome. Just because we are able to demonstrate that a program was effective in terms of certain outcomes, we cannot prove that the results were due to the program per se. Positive results, based on one program, also cannot be generalized to other programs in other communities (Cronbach, 1974). One example is the recent publication of the much sought-after, positive results of infant and preschool intervention projects (Lazar, et al., 1977). The findings were based on 14 experimental programs with "deliberate cognitive curricula" (Lazar et al., 1977, p. 2). Many of the programs were university-based and small in scale, and received closer control and supervision than most early childhood programs.

Another challenge to generalizability of program effects has been raised by House and his colleagues in their critique of the Abt Associates' evaluation of Planned Variation Follow Through. House, Glass, McLean, and Walker (1978) stressed the importance of "local individuality" (p. 474), meaning that models that worked well in one community worked poorly in another. They argued that local setting variables (individual teachers, schools, neighborhoods, homes) had more effect on achievement than did the labelled models.

These reservations will raise the wrath of some of my colleagues who will undoubtedly accuse me of joining the ranks of Jensen, Herrnstein, Eyenck, and others who do not have much faith in the results of Headstart and other early childhood programs. I will probably be accused of making statements detrimental to the future funding of early childhood programs just when we appear to have evidence of long-term effects. Hence, these comments are made with uncertainty and regret. There is, however, a hopeful side to my observations. A developmental perspective to evaluation with its dimensions of reciprocal causality and organismic complexity discussed earlier suggests that the search for long-term effects of programs is illusionary. In the linear-unidirectional causation model, the preschool intervention program, often conceived of as the "treatment," is assumed to affect a child's development as a significant main effect and have persistent effects later in the child's life. Koocher and Broskowski (1977) have labeled this mode of thinking as "the single-input fallacy" (p. 584). Assessment of program impact must take into account multiple and competing influences on the child during the program and in the period of time after the program ends and the follow-up study begins. When the child is in the program and in the intervening years, she is exposed to multiple socializing influences, many of which we are not able to (and possibly should not) control. Concerns over "fadeout" or maintenance of "gains" are considerably diminished.

I am aware that committed advocates of early childhood programs strongly believe that they must demonstrate the long-term effectiveness of preschool intervention (often without knowing how programs function and why effects are observed) before public monies will flow again. But I would argue as advocates for young children, we have gotten ourselves

in a peculiar situation. We are not in a position from a research and an evaluation stance to demonstrate the long-term effectiveness of any early childhood program. The presumed demands of the policy makers outstrip the state of the art in methods, instruments, and resources (Zimiles, 1977). Yet we persist in doing so, because we think that is the kind of information policy makers want.[5] Perhaps this question reflects my limited experience and naivete, but why do we try to build up expectations that we know will go unmet? Why not say we cannot at present demonstrate effectiveness, and that we should have programs for children for certain desirable social goals and values, and not because the evaluations say we should? Perhaps part of the answer lies in the tendency of American social scientists to believe that the demonstration of the effectiveness of social programs is imminently possible and that policy decisions will be made in a rational, systematic manner (Sarason, 1978).

We should take cognizance of the conclusions of a number of contributors to C. Weiss' (1977) volume Using Social Research in Public Policy Making that evaluation findings remain only one source of information for decision makers. Ethical, legal, economic, and political considerations are also important factors in decision making. Datta'a (1976) detailed account of the impact of the Westinghouse/Ohio evaluation on the development of Head Start confirms the importance of these factors and their complex, often unpredictable interrelationships.

Preliminaries to a developmental perspective to evaluation.

Ross and Cronbach (1976) state that "what epistemological stance, what view of the political system, or what assumptions about the purposes of evaluation lead to each of the divergent positions taken by evaluators,

researchers, and decision makers . . ." (p. 19). The epistemological stance has already been discussed in the presentations of alternative models of development. What is being proposed here is a rationale for working toward a developmental perspective, which represents a closer match between evaluation strategies and the phenomena under study. I am not yet ready to present this perspective in great detail at this point, but will share some of my thoughts-in-formation. In this section, views of the political system and assumptions about the purposes of evaluation enter into a rationale for a developmental perspective to evaluation of early childhood programs. In all my considerations, I have been guided by Albert Einsein's statement: "A perfection of means and confusion of aims seems to be our main problem."

What Are the Purposes of an Early Childhood Evaluation?

The question of the purposes of an evaluation is critical because purposes shape the design, instrumentation, and interpretation of an evaluation (Messick, 1975). If, for example, the purpose of an early childhood program is to increase school achievement and IQ test scores, then an evaluation may be designed to assess them easily. However, a number of early educators have argued that increases in school achievement and in performance on IQ tests are not the purpose of early childhood programs (Kamii, 1973), particularly of programs that are not aimed toward achievement in the school-sense (Kohlberg & Mayer, 1973). Hence, standardized tests are not appropriate.

From the models of development presented in the previous section, it follows that a developmental perspective to evaluation is not primarily concerned with achievement data. On the contrary, the purposes of evaluation of early childhood programs are directed toward involving the

participants, including the evaluators, in understanding program development, adult (staff and parent) development, and child development. Evaluation is directed toward enabling all involved in the program to reflect critically about what is happening (Carini, 1975; Eisner, 1977; Cronbach, 1974; Partlett & Hamilton, 1976). Thus evaluation is geared toward providing feedback about the program to the staff and other audiences in terms that are comprehensible and which lead to more sound practices.[6] The most important criterion against which an evaluation is judged is its utility, or the extent to which the evaluation results in program improvement and child development.

A related criterion for an evaluation is the degree to which it enhances the development of all participants and their dignity as human beings (Sjoberg, 1975; Tyler, 1978). Thus, considerable attention should be focused on how evaluation is actually practiced and how it affects the participants' self-esteem (Report of the Task Force on Testing and Assessment of Children, 1976; Tyler, 1978).

A developmental perspective to the evaluation of programs is compatible with more formative modes of evaluation and with an "extended" view of evaluation described by Ross and Cronbach (1976). Evaluation is seen as a continuing part of program management and planning. This view is also consistent with Public Law (PL) 94-63 of the Community Mental Health Centers (CMHC) Amendments of 1975 (Davis, Windle, & Sharfstein, 1977; Guidelines for Program Evaluation in CMHCs, 1977) which obligates federally funded community mental health centers to conduct program evaluations on an ongoing basis to improve services and to be more responsive to clients. The evaluator is part of the program; she studies what was delivered and how people interacted during the program. The

evaluator functions as a naturalistic observer whose inquiry grows out of his/her observations. These strategies are also compatible with a qualitative approach to evaluation (Eisner, 1977; Hamilton et al., 1977; Stake, 1967; Wilson, 1977; Willis, 1978; Wolf & Tymitz, 1977; Campbell, Note 1).

A developmental perspective to the the evaluation of programs. In evaluations based on the experimental paradigm, programs are "treatments" which are assumed to be similar across sites and which can be replicated in other sites. This conceptualization of programs is based on the ideal of generalizable knowledge (Cronbach, 1975). However, large-scale evaluations of early childhood programs indicate that the assumption that programs can be treated as a set of unitary variables is questionable. House, Glass, McLean and Walker (1978) point to local setting variables as more important in determining achievement scores than the labelled program models in Project Follow Through.

An alternative conceptualization is programs-as-cultural systems (in contrast to programs-as-treatments) with histories, traditions, and values. The anthropologist Clifford Geertz (1973) has aptly expressed the distinction I am making: "The concept of culture I espouse . . . is essentially a semiotic one. Believing, with Max Weber, that man is an animal suspended in webs of significance he himself has spun, I take culture to be those webs, and the analysis of it to be therefore not an experimental science in search of law but an interpretative one in search of meaning" (p. 5). Of central interest in evaluations of programs would be how children and adults structure meaning out of their experiences and create the daily routine of program life (see, for example, Mehan, 1978). It is understood that in any program there will be multiple

interpretations of what the program is (Hamilton, et al., 1977). Methods for studying programs-as-cultures may include case studies (Hamilton, et al., 1977; MacDonald & Walker, 1975; Stake, 1978), ethnographic approaches (Mehan, 1978; Mehan, et al., Note 3); documentation (Carini, 1975; Hamilton & Partlett, 1976), among others in a rich, emerging area (Willis, 1978).

A developmental perspective is also concerned with the study of programs and classroom environments in terms of their change over time. Guttentag (1977) articulated this idea: "Given the nature of most social programs, it is not realistic to presume that they can be kept the same over time . . . . When evaluators act as though the program was indeed unchanging over time, it is nearly always a fiction. Results based on such seemingly 'unchanging' programs are hard to believe, and may lead to false inferences" (p. 19).

Most studies of classroom environments assume that they do not change over time, such as the school year. However, people who are involved with early childhood programs know that there are forces for change--both improvement and regression. Programs rarely remain static entities. For example, in Weikart's (1972) Ypsilanti Project, there was continuous opportunity for improvement and change in what teachers were doing in their classrooms. A developmental perspective is sensitive to these changes within programs and concerned with documenting and describing them in a way which can be related to the program's influence on adults and children. Evaluation is part of a dynamically evolving program.

A developmental perspective to adult development. Evaluation of early childhood programs have been almost exclusively focused on children. Understanding staff relationships and parent and staff development is

important to the provision of early childhood programs and should be systematically examined in future evaluations.

Katz (1971; 1973; 1977) has outlined some problems in the sociology of early childhood education which take into account the social and political context in which early childhood educators work. The effects of mandated, often shifting program requirements, and the conditions of the work place, particularly the special problems involved in working with young children (Katz, 1971; 1977) should be examined in evaluations. Developmental stages of preschool teachers (Katz, 1972) might be applied to looking at staff development in programs. Weikart (1972) argued that teacher motivation, supervision, and resourcefulness appear to be particularly important in program operations. He argued that any project must have an effective staff model which includes high intensity of planning and supervision.

Another potential area for evaluation is how early educators accept and use developmental theories and research, and curriculum and instructional materials. This kind of examination challenges the taken-for-granted assumption that research can provide suggestions how to teach by the translation of child development research into curriculum (Katz, 1973). Jackson and Kieslar (1977) have effectively challenged this linear approach to the relation between research and practice. They point out that we know very little about how theory and research actually become part of educational practice. Of particular interest is how theory and research matches or fits into the "vision of reality" (Jackson & Kieslar, 1977) of the teachers. Inasmuch as the educators' view of reality is linked to possible adaptive consequences of working with young children (Katz, 1971), research into the belief and culture of

early educators is important in understanding how programs develop over time, and is an important component of qualitative evaluations of educational innovations (Magoon, 1977).

A developmental perspective to evaluation would also focus seriously on understanding parental and family development in early childhood programs. Reviews of the effects of intervention programs identify the importance of the mother's teaching style and attitudes, and her role in "maintaining gains" of the programs (Bronfenbrenner, 1973; Shipman et al., 1976). Sigel (1972) pointed to the importance of considering changes in the child as a result of the program in relation to changes in the behavior and attitudes of the parents. Mothers are reported to go back to work and school and to be involved in community life as a result of their children's participation in the program. Yet our knowledge of the relationships between early childhood programs and parental and family development remains supplemental in most evaluations, often treated as anecdotal asides (see, however, Falender & Heber, 1975; Slaughter, Note 7).

Evaluations in this area should not focus solely on parental behavior vis-a-vis the child, but also on parents as persons with their own needs and interests. Research on parenthood as a developmental stage (Benedek, 1959; Leifer, 1977) provides some concepts for examining the adaptations parents may make to their or their child's involvement in an early childhood program. Brown's (1975) work on family dialectics in clinical situations is also suggestive for closer examinations of family development when parents participate in programs designed for them and/or for their children.

A developmental perspective to assessing children. It is striking that in the many evaluations of early childhood programs, many of which are in operation for the same children over a number of years, that we know very little about how individual children develop in these programs.[7] Although staff talk a great deal about how individual children develop in their programs, we have few descriptive, longitudinal records of assessment of change over age within individuals. There are two reasons why more attention should be placed on the child as the unit of analysis. If one of the goals of early childhood programs is to enhance the child's life chances, then the individual is the appropriate unit of analysis. Furthermore, in longitudinal studies, the only unit that has continuity over time is the individual child (Haney, Note 2). There is a great need for individual developmental histories of children in programs like those constructed in the Berkeley Growth Studies (Jones, Bayley, Macfarlane, & Honzik, 1971) and Murphy's (1962) studies of the Topeka children. Shipman et al. (Note 6) intensive case studies from the ETS-Head Start longitudinal sample augur well for the future.

The importance of longitudinal studies has been identified by many workers in the field (Lazar et al., 1977; Shipman, 1976; Sigel, Secrist, & Forman, 1972; Shipman, et al., Note 6). However, fundamental problems of how development and change will be conceptualized in these longitudinal studies needs more systematic attention. In her case studies, for example, Shipman et al. (Note 6) pointed to the inappropriateness of linear analytic methods given the unevenness of developmental change. Sigel, Secrist, and Forman (1972) described problems in longitudinal studies when conducted in intervention programs serving low-income populations. The most serious problem, however, is the fact that

longitudinal studies are relatively rare and remain an underdeveloped field both in conceptualization and methodology (McCall, 1977; Wohlhill, 1973).

Whom Shall the Evaluation Serve?

Related to the question of the purposes of an evaluation is the issue of whom the evaluation should serve. Sjoberg (1975) charged that evaluators usually align themselves with the powerful or dominant groups in the system and accept these groups' definitions of program goals and desired outcomes. Thus evaluation serves several functions including reform, manipulation, and sustaining power or structural relations. Sjoberg argues for a countersystem role in which the evaluator works with the less powerful in the system. In a similar vein, House argues for justice-as-fairness as an important standard for evaluation. "By the second principle (of justice), social and economic inequalities must benefit the least advantaged in the long run. The educationally least advantaged within most settings are the children first and the teachers second. The evaluator should strive to present their views and perspectives" (p. 97).

In early childhood programs, staff and parents have a particular stake in the program and its services. They are the individuals who are actively involved in implementing the program. If evaluation is to have utility, these individuals must be participants in the process.[8] An example of how evaluation purposes change depending upon the role and perspectives of the groups involved is illustrated in a report on evaluation in Head Start programs written by a group composed largely of staff and parents (Report of the Task Force on Testing and Assessment of Children, 1976). The Task Force advocated procedures "which recognize

and support individual rights, facilitate growth and enhance dignity" (p. 17). Its "Statement of Rights" for children, parents, and staff can be found in Appendix I.

It should be clear by now that my bias in arguing for a developmental perspective is that the sole purpose of evaluation is not oriented toward judgment of the effectiveness or efficiency of a program or toward the perceived needs of public policy makers for quantitative, "hard" data based on large samples of children. In fact, serious challenges have been raised about the evaluator's assumption that evaluations provide information for decision makers (Weiss, 1977; Wise, Note 9). Others have pointed to the limitations of large-scale survey research for policy making (Mehan, 1978). Findings from large-scale studies are probabilistic and do not apply to particular programs. They rarely reveal much about the processes which create and maintain programs, and hence are limited in identifying specific actions for change. Finally, since the findings are abstract rather than concrete, motivating staff concern for improvement is understandably difficult.

The purposes of evaluation are based on House's (1976) notions of justice in evaluation and are oriented toward program development and serving needs of the staff and children, not primarily the bureaucratic and funding agencies. These purposes also reflect changing conceptions and diversification of evaluation activity in the '70s toward program evaluation and qualitative approaches (Hamilton et al., 1977; Tikunoff & Ward, 1977; Willis, 1978). What I am arguing for, then, is that there must be cultivation of alternatives to what presently exists in the evaluation of early childhood programs, recognition of the importance of values in choosing methods, and care in conducting systematic, disciplined

inquiry. We have many more options in evaluating early childhood programs than we are currently utilizing (Perrone, Cohen, & Martin, Note 5). Explorations of alternatives will allow us to see new problems, invent new methods, and to understand our programs in ways that better fit our experiences of the complexities of program life and of adult and child development.

Constraints on a Developmental Perspective

An argument for a developmental perspective to evaluation of early childhood programs has been presented. There are, however, a number of constraints on the further development of this perspective. One major constraint is that research in child development has itself neglected a developmental orientation (McCall, 1977; Wohlhill, 1973). Furthermore, there has been an emphasis on concepts of stability, continuity, and equilibrium over those of instability, discontinuity, and change in longitudinal research (Riegel, 1972; 1976; McCall, Eichorn, & Hogarty, 1977; Wohlhill, 1973). Thus, a conceptual and methodological base for a developmental perspective is underdeveloped.

Another critical constraint is that a developmental perspective to evaluation represents assumptions and ideologies which are gaining ascendance,[9] but which are still overshadowed by what have been variously called quantitative or experimental approaches to evaluation. This paper posited earlier that paradigms of evaluation represent views of reality and the selection of one over another is a matter of values. According to Kuhn's notion of the paradigm, a paradigm gains power and dominance because most of the researchers, through socialization into their discipline, agree to believe and trust it.

Levine (1974), among others, observes that science is socially constructed by a community: "Science is what scientists feel comfortable in recommending to others as principles through which the world may be manipulated, predicted, or understood. Science is what scientists say it is at any given point in time. Science is what scientists feel comfortable in writing about in articles and in textbooks" (p. 669).

Developmental perspectives to evaluation which are akin to qualitative, ethnographic, and naturalistic approaches to evaluation are weak in comparison to the dominant model, lack a strong and sizable community of committed workers, and suffer from the lack of legitimation. Strategies for creating parity and tolerance are needed simultaneously with further work on a developmental perspective. Messick (1975) offers Churchman's proposals for the study of systems of inquiry that can aid in the exposure of the implicit value assumptions in research strategies. Levine (1974) suggests an adversary model. Meanwhile, one constraint on the further development of a developmental perspective is the dominance of the experimental and psychometric models.

But there is an even more troublesome constraint related to complex issues in the ethics of evaluation studies. Evaluations that aim at the intensive documentation of program life and at the interpretative understandings of the participants may be too revealing. These evaluations may unmask the protective myths surrounding early childhood programs. Quantitative methods, by their very nature, do not have the potential of such illumination. Can we know too much? For what ends will knowledge from these studies be used? Indeed, we are traveling into dangerous uncharted territory. The lack of guidelines for evaluators who engage in qualitative approaches is a central concern of individuals who have used them (MacDonald & Walker, 1975).

Given these very significant and powerful constraints, individuals who are involved in the evaluation of early childhood programs face a major reconsideration of their purposes and methods. Moving toward a developmental perspective calls for a reorientation in many of the ways evaluation is currently practiced and supported. In this paper, I have aspired to lay the groundwork for an evaluation model that is grounded in knowledge of the traditions, problems, and practicing realities of early childhood programs and which in all its considerations places ideas of development as the center of inquiry. In reflecting on the need for and future work on a developmental perspective, I am encouraged by John Tukey's observation: "Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong questions, which can always be made precise" (cited in Rose, 1977, p. 23).

ACKNOWLEDGEMENTS

This paper brings together ideas in evaluation, developmental theory and research, and early childhood programs, and reflects my values, training, and experience. The act of pointing to the source of one's ideas is complex, but I would like to acknowledge some especially formative individuals: Lee J. Cronbach, John C. Glidewell, Louis L. Knowles, Beryl Scoles, Louise L. Tyler, and Docia Zavitkovsky. I am grateful to Louise Tyler for her critical comments on an earlier draft of this paper. Lucinda Bernheimer and Arline Dillman provided editorial assistance. Of course, I assume sole responsibility for the constructions and interpretations I have made in this integrative review paper. Work on this paper was supported in part by a Spencer Fellowship from the National Academy of Education.

Reference Notes

1. Campbell, D. Qualitative knowing in action research. Paper presented at the meeting of the American Psychological Association, New Orleans, September 1974.
2. Haney, W. Units of analysis issues in the evaluation of Project Follow Through. Unpublished manuscript, Huron Institute, 1974.
3. Mehan, H., Cazden, C. B., Coles, L., Fisher, S., & Maroules, N. The social organization of classroom lessons. (Center for Human Information Processing Report No. 67). University of California, San Diego, 1976.
4. Patton, M. Q. Alternative evaluation paradigm. Grand Forks, N.D.: Center for Teaching and Learning, University of North Dakota, 1975.
5. Perrone, V., Cohen, M. D., & Martin, L. P. (Eds.) Testing and evaluation. New views. Washington, D.C.: Association of Childhood Education International, 1975.
6. Shipman, V. C., et al. Notable early characteristics of high and low achieving Black low-SES children. Princeton, N.J.: Educational Testing Service, 1976.
7. Slaughter, D. T. Parent education for low income black families. Presented at the General Mills American Family Forum on "Parenting--The Crucial Years," Washington, D.C.: October 17-18, 1977.
8. Takanishi, R., & Spitzer, S. Children's perceptions of human resources in team teaching classrooms. A cross-sectional developmental study. Paper presented at the meeting of the American Educational Research Association, Toronto, Canada, March 1978.

9. Wise, R. I. What we know about the decision-maker and decision settings. Paper presented at the American Educational Research Association, Toronto, Canada, March 1978.

10. Zigler, E. The effectiveness of Head Start: Another look. Paper presented at the meeting of the American Psychological Association, San Francisco, August 1977.

References

Almy, M. The early educator at work. New York: McGraw-Hill, 1975.

Ball, S., & Bogatz, G. The first year of Sesame Street: An evaluation. Princeton, N.J.: Educational Testing Service, 1970.

Benedek, T. Parenthood as a developmental phase. Journal of the American Psychoanalytic Association, 1959, 7, 389-417.

Bronfenbrenner, U. Is early intervention effective? Washington, D.C.: U.S. Office of Child Development, 1973.

Brown, L. K. Familial dialectics in a clinical context. Human Development, 1975, 18, 223-238.

Carini, P. F. Observation and description. Grand Forks, N.D.: University of North Dakota Press, 1975.

Cazden, C. The neglected situation in child language research and education. In F. Williams (Ed.), Language and poverty: Perspectives on a theme. Chicago: Martham, 1970.

Cicourel, A. V., Jennings, K. H., Jennings, S. H. M., Leiter, K, C. W., MacKay, R., Mehan, H., & Roth, D. R. Language use and school performance. New York: Academic Press, 1974.

Cook, T., Appleton, R., Conner, R., Schaffer, G., Tamkin, G., & Weber, S. Sesame Street revisited. New York: Russell Sage Foundation, 1975.

Cronbach, L. J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 115-127.

Datta, L. The impact of the Westinghouse/Ohio evaluation on the development of Project Head Start: An examination of the immediate and longer-term effects and how they came about. In C. C. Abt (Ed.), The evaluation of social program. Beverly Hills, CA.: Sage, 1976.

Davis, H. R., Windle, C., & Sharfstein, S. S. Developing guidelines for program evaluation capability in community mental health centers. Evaluation, 1977, 4, 25-29.

DeVries, R. Relationships among Piagetian, IQ, and achievement assessments. Child Development, 1974, 45, 746-56.

DeVries, R., & Kohlberg, L. Relations between Piagetian and psychometric assessments of intelligence. In L. G. Katz (Ed.), Current topics in early childhood education (Vol. I). Norwood, N.J.: ABLEX, 1977.

Eisner, E. W. On the uses of educational connoisseurship and criticism for evaluating classroom life. Teachers College Record, 1977, 78, 345-58.

Falender, C. A., & Heber, R. Mother-child interaction and participation in a longitudinal intervention program. Developmental Psychology, 1975, 11, 830-36.

Frank, L. K. Evaluation of educational programs. Young Children, 1968, 24, 167-74.

Geertz, C. The interpretation of cultures. New York: Basic Books, 1973.

Glick, J. Some problems in the evaluation of preschool intervention programs. In R. D. Hess and R. M. Bear (Eds.), Early education. Chicago: Aldine, 1968.

Guba, E. G. The failure of educational evaluation. Educational Technology, 1969, 9(5), 29-38.

Guidelines for program evaluation in CMHCs. Evaluation, 1977, 4, 30-34.

Guttentag, M. On quantified sachel: A reply to Apsler. Evaluation, 1977, 4, 18-20.

Hamilton, D., Jenkins, D., King, C., MacDonald, B., & Parlett, M. Beyond the numbers game: A reader in educational evaluation. Berkeley: McCutchan, 1977.

Harre, R. The conditions for a social psychology of childhood. In M. P. M. Richards (Ed.), The integation of a child into a social world. New York: Cambridge University Press, 1974.

House, E. R. (Ed.) School evaluation: The politics and process. Berkeley, Calif.: McCutchan, 1973.

House, E. R. Justice in evaluation. In G. V. Glass (Ed.), Evaluation studies review annual (Vol. 1). Beverly Hills, Calif.: Sage, 1976.

House, E. R., Glass, G. V., McLean, L. D., & Walker, D. F. Critiquing a Follow Through evaluation. Phi Delta Kappan, 1978, 59, 473-474.

Jensen, A. How much can we boost IQ and school achievement? Harvard Educational Review, 1969, 39, 1-123.

Jones, M. C., Bayley, N., Macfarlane, J., & Honzik, M. The course of human development. Waltham, Mass.: Xerox College Publishing, 1971.

Kamii, C. K. Evaluation of learning in preschool education: Socio-emotional, perceptual-motor, cogniive development. In B. S. Bloom, J. T. Hastings, and G. F. Madaus, Handbook on formative and summative evaluation. New York: McGraw-Hill, 1971.

Kamii, C. One intelligence indivisible. Young Children, 1975, 30, 228-38.

Kamii, C., & Elliott, D. L. Evaluation of evaluations. Educational Leadership, 1971, 28(8), 827-31.

Katz, L. G. Sentimentality in preschool teachers: Some possible interpretations. Peabody Journal of Education, 1971, 48, 96-105.

Katz, L. G. Developmental stages of preschool teachers. The Elementary School Journal, 1972, 23, 50-54.

Katz, L. G. Where is early childhood education going? Theory into Practice, 1973, 12, 137-42.

Katz, L. G. Early childhood programs and ideological disputes. Educational Forum, 1975, 10, 267-71.

Katz, L. G. Teachers in preschools: Problems and prospects. International Journal of Early Childhood, 1977, 9, 111-23.

Kilmer, S., & Weinberg, R. The nature of young children and the state of early education: Reflections from the Minnesota Roundtable. Young Children, 1974, 30(1), 60-67.

Kohlberg, L., & Mayer, R. Development as the aim of education. Harvard Educational Review, 1972, 42, 449-96.

Koocher, G., & Broskowski, A. Issues in the evaluation of mental health services for children. Professional Psychology, 1977, 8, 583-92.

Kuhn, T. The structure of scientific revolutions. Chicago: University of Chicago Press, 1962.

Kvale, S. The technological paradigm of psychological research. Journal of Phenomenological Psychology, 1973, 3, 143-59.

Kvale, S. The psychology of learning as ideology and technology. Behaviorism, 1976, 4, 97-116.

Lazar, I. et al. The persistence of preschool effects. A long-term follow-up of 14 infant and preschool experiments. Washington, D.C.: U.S. Government Printing Office, 1977.

Leifer, M. Psychological changes accompanying pregnancy and motherhood. Genetic Psychology Monographs, 1977, 95, 55-96.

Levine, M. Scientific method and the adversary model: Some preliminary thoughts. American Psychologist, 1974, 29, 661-77.

Looft, W. R. Socialization and personality throughout the life span: An examination of contemporary psychological approaches. In P. B. Baltes and K. W. Schaie (Eds.), Life-span developmental psychology. Personality and socialization. New York: Academic Press, 1973.

MacDonald, B., & Walker, R. Case study and the social philosophy of educational research. Cambridge Journal of Education, 1975, 5.

Magoon, A. J. Constructivist research. Review of Education, 1977, 49, 651-94.

McCall, R. B. Challenges to a science of developmental psychology. Child Development, 1977, 48, 333-44.

McCall, R. B., Eichorn, D. H., & Hogarty, P. S. Transitions in early mental development with commentary by I. C. Uzgiris and by E. S. Schaefer, with reply by the authors. Monographs of the Society for Research in Child Development, 1977, 42 (3, Serial No. 171).

McClelland, D. C. Testing for competence rather than for "intelligence." American Psychologist, 1973, 28, 1-14.

Mehan, H. Structuring school structure. Harvard Educational Review, 1978, 48, 32-64.

Merleau-Ponty, M. Phenomenology of perception (trans. C. Smith). London: Routledge and Kegan Paul, 1962.

Messick, S. The standard problem. Meaning and values in measurement and evaluation. American Psychologist, 1975, 30, 955-66.

Murphy, L. B. The widening world of childhood. New York: Basic Books, 1962.

Murphy, L. B. The strangehold of norms on the individual child. Childhood Education, 1973, 49, 343-49.

Overton, W. F., & Reese, H. W. Models of development: Methodological implications. In J. R. Nesselroade and H. W. Reese (Eds.), Life-span developmental psychology: Methodological issues. New York: Academic Press, 1973.

Partlett, M., & Hamilton, D. Evaluation as illumination: A new approach to the study of innovative programs. In G. V. Glass (Ed.), Evaluation studies review annual (Vol. 1). Beverly Hills: Sage, 1976.

Petrie, H. G. Theories are tested by observing the facts: Or are they? In L. G. Thomas (Ed.), Philosophical redirection of educational research. Seventy-first yearbook of the National Society for the Study of Education, Part 1. Chicago: University of Chicago Press, 1972.

Pinar, W. F. Currere: A case study. In G. Willis (Ed.), Qualitative evaluation: Concepts and cases in curriculum criticism. Berkeley, CA.: McCutchan, 1978.

Reese, H. W., & Overton, W. F. Models of development and theories of development. In L. R. Goulet and P. B. Baltes (Eds.), Life-span developmental psychology. Research and theory. New York: Academic Press, 1970.

Riegel, K. F. Influence of economic and political ideologies on the development of developmental psychology. Psychological Bulletin, 1972, 78, 129-41.

Riegel, K. F. From traits and equilibrium to developmental dialectics. In W. J. Arnold (Ed.), Nebraska Symposium on Motivation, 1975. Lincoln: University of Nebraska Press, 1976.

Rose, R. Disciplined research and undisciplined problems. In C. H. Weiss (Ed.), Using social research in public policy making. Lexington, Mass.: Lexington Books, 1977.

Ross, L., & Cronbach, L. Handbook of evaluation research: Essay review. Educational Researcher, 1976, 5(10), 9-10.

Sameroff, A. J. Early influences on development: Fact or fancy? Merrill-Palmer Quarterly, 1975, 21, 267-294.

Sarason, S. B. The nature of problem solving in social action. American Psychologist, 1978, 33, 370-380.

Shapiro, E. Educational evaluation: Rethinking the criteria of competence. School Review, 1973, 81, 523-50.

Sigel, I. Developmental theory: Its place and relevance in early intervention programs. Young Children, 1972, 27, 364-72.

Sigel, I. E., Secrist, A., & Forman, G. Psycho-educational intervention beginning at age two: Reflections and outcomes. In J. C. Stanley (Ed.), Compensatory education for children ages two to eight: Recent studies of educational intervention. Baltimore: Johns Hopkins University Press, 1972.

Sjoberg, G. Politics, ethics, and evaluation research. In M. Guttentag and E. L. Struening (Eds.), Handbook of evaluation research (Vol. 2). Beverly Hills: Sage Publications, 1975.

Skinner, Q. The flight from positivism (Review of R. J. Bernstein's The restructuring of social and political theory). New York Review of Books, 1978, 25, 26-28.

Stake, R. E. The case study method in social inquiry. Educational Researcher, 1978, 7, 5-8.

Stake, R. E. The countenance of educational evaluation. Teachers College Record, 1967, 68, 523-40.

Strike, K. A. Explaining and understanding: The impact of science on our concept of man. In L. E. Thomas (Ed.), Philosophical redirection of educational research. The seventy-first yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press, 1972.

Tikunoff, W. J., & Ward, B. A. (Eds.) Exploring qualitative/quantitative research methodologies in education. Anthropology and Education Quarterly, 1977, 8.

Tyler, L. L. Curriculum evaluation and persons. Educational Leadership, 1978, 35, 275-79.

Weikart, D. P. Relationship of curriculum, teaching, and learning in preschool education. In J. C. Stanley (Ed.), Preschool programs for the disadvantaged. Baltimore: Johns Hopkins, 1972.

Weiss, C. Using social research in public policy making. Lexington, Mass.: D. C. Heath and Co., 1977.

Willis, G. Qualitative evaluation. Concepts and cases in curriculum criticism. Berkeley: McCutchan, 1978.

Wilson, S. The use of ethnographic techniques in educational research. Review of Educational Research, 1977, 47, 245-65.

Wohlwill, J. The study of behavioral development. New York: Academic Press, 1973.

Wolf, R. L., & Tymitz, B. Ethnography and reading: Matching inquiry mode to process. Reading Research Quarterly, 1977, 12.

Zigler, E., & Butterfield, E. C. Motivational aspects of changes in IQ test performance of culturally deprived nursery school children. Child Development, 1968, 39, 1-14.

Zimiles, H. A radical and regressive solution to the problem of evaluation. In L. G. Katz (Ed.), Current topics in early childhood education, (Vol. 1), Norwood, N.J.: ABLEX, 1977.

Footnotes

[1]A version of this paper was presented at the annual meeting of the American Educational Research Association, Toronto, Canada, March 28, 1978.

[2]Economic and political ideologies have influenced the concept of development in American developmental psychology (Riegel, 1972). Riegel contended that quantitative models are associated with the "capitalistic" orientation of the Anglo-American countries, and represent continuous growth models in which all individuals are evaluated against single standards. The "mercentalistic-socialistic" orientation of the European continent is represented by qualitative models which focus on the organization and structure of experiences and which evaluate individuals according to multiple standards. Detailed analyses of the ideological and philosophical underpinnings of models of development have been described elsewhere (Riegel, 1972; Kvale, 1973, 1976) and will not be covered here.

[3]Pinar's (1978) essay on currere, the analysis of the individual's lived experience of a curriculum, richly expands on the ideas expressed here.

[4]For an excellent and detailed discussion of this research, see Mehan (1978), pp. 49-56.

[5]A problem worthy of careful consideration is what kinds of information policy makers want about early childhood programs. The focus on child outcomes, specifically measured by intelligence and achievement tests, is itself an interesting social fact. Most of the arguments that Head Start "does not work" (which are generalized to

other early intervention programs) are based solely on cognitive outcomes. However, there is considerable agreement that Head Start "works" in areas related to health screening and treatment, nutrition, and parent and community participation. Head Start serves as a potential model for the coordination of child and family services. This coordination of services was, in fact, a major thrust of its original mission (Datta, 1976).

[6]For an extended discussion of this point, see Stake (1978).

[7]Existing laws (PL 94-142 and PL 93-380) require that certain programs determine their impact on individual children. According to PL 94-142, teachers must evaluate the child's progress according to the Individualized Educational Plan (IEP).

[8]I recognize that young children must also be considered as participants in evaluations, but am not ready to deal with this problem in depth at this time.

[9]For an excellent, integrative review of declining faith in mainstream social science, see Skinner (1978).

APPENDIX I

V. CONCLUSION

This Task Force has considered a range of issues regarding assessment: need, process, use, and impact, in the body of this paper. In reviewing the need for assessment, the Task Force recognizes that assessment of specific developmental processes in young children is an established and necessary practice. The challenge for those who work with children is to utilize assessment methods which recognize and support individual rights, facilitate growth, and enhance dignity.

After exploring the process and use of assessment, the Task Force has agreed that there is a wide variation in the quality of assessment methods available for use with young children. Therefore, care must be taken in the selection of instruments and assessors, in order to provide useful information for planning educational opportunities.

Finally, the Task Force recognizes that the effect of the assessment process on a child's family is a critically important factor, and must be taken into consideration at each step in that process. The Task Force members have seriously studied these issues, and as a result of their efforts, have developed a Statement of Rights for children, parents, and staff as the blueprint for implementation in any program serving young children and their families. Although parents and staff fulfill the role of advocates for children's rights, they also have rights in this regard.

This Task Force recommends that all programs for young children adopt the following Statement of Rights:

Rights of the Child

A child has the right to be different, and to be accepted as such. Differences in individual children should be approached in a positive, meaningful way so they may function to their fullest capacity in a pluralistic society.

A child has the right to be assessed, and as a result through an assessment, to be provided with a quality developmental program.

A child has the right to be tested under optimal conditions in a non-threatening environment by a person sensitive to children.

SOURCE: Report of the Task Force on Testing and Assessment of Children (Submitted to the Office of Child Development). San Francisco: Development Associates, 1976.

A child has the right to be assessed with a nonbiased instrument by a person who speaks the language in which the child is most fluent.

A child who scores differently from the norm on any given test or assessment has the right not to be labeled.

A child has the right to have his/her observable behavior recorded in functionally descriptive terms rather than in generalized terminology (labeling).

A child has the right to have the results of assessments kept confidential, and the records kept in a locked file. A child's records shall only be made accessible to the child's parents and other authorized persons.

Rights of Parents

Parents have the right that their child receive an overall assessment which includes information obtained from health examinations, classroom observations, parent conferences, and home visits.

Parents have the right to be informed of the purpose of the assessment, and of the instrument(s) to be used in assessing their child.

Parents have the right to give or withhold permission to have their child assessed, and to challenge the content of written records.

Parents have the right to be involved in the total assessment process.

Parents have a right to give input into the overall assessment of their child and provide the person(s) doing the assessment with their views and observations of the behavior, development, and activities of their child.

Parents have the right to be informed about the assessment results, and to have conferences with appropriate staff regarding interpretation of the results and for program planning.

Parents have a right, at their request, to a written summary of a conference following the assessment.

Parents have the right to confidentiality of all information obtained by the assessor.

Parents have rights regarding the releasing of assessment results to other agencies or public schools. Unless they have given permission, the assessment information shall not be forwarded.

Parents have the right to be treated with consideration and sensitivity regarding the psychological and social impact effects that assessment results can have on the family unit.

Parents have the right to receive all supportive services necessary to effectively implement assessment results.

Rights of Staff

The teaching staff has the right to not be overburdened by the assessment process and recordkeeping to the extent that it interferes with their primary teaching function. It should be recognized, however, that making and recording observations of children are an integral part of teaching.

Staff has the right to give input into the selection and appropriateness of assessment instruments (for both group and individual use), and to understand the relevance of the assessment results to the overall program.

Staff has the right to request and receive proper instruction and training in the use of the assessment instrument and consultation in the interpretation of the results.

Staff has the right to receive information, in advance of the time, methods, persons involved, and space to be used when assessment activities will be conducted.

Staff has the right to object to the use of a specific instrument which in their professional opinion, goes contrary to good child development principles.

Staff has the right to receive timely feedback after the assessment has been completed.

MS11:A